Subterranean Biology 38: 47–63 (2021)
doi: 10.3897/subtbiol.38.60865
https://subtbiol.pensoft.net

RESEARCH ARTICLE

A peer-reviewed open-access journal
Subterranean Biology
Published by The International Society for Subterranean Biology

The Pennsylvania grotto sculpin: population genetics

Luis Espinasa[1], Drake M. Smith[2], Julianna M. Lindquist[1]

1 *School of Science, Marist College, 3399 North Rd, Poughkeepsie, New York 12601, USA* **2** *Texas Tech University, 2500 Broadway Lubbock, Texas 79409*

Corresponding author: Luis Espinasa (luis.espinasa@marist.edu)

Academic editor: M.L. Niemiller | Received 17 November 2020 | Accepted 16 March 2021 | Published 12 April 2021

http://zoobank.org/D7171664-2C1A-43A4-8EBC-6FD57374F654

Citation: Espinasa L, Smith DM, Lindquist JM (2021) The Pennsylvania grotto sculpin: population genetics. Subterranean Biology 38: 47–63. https://doi.org/10.3897/subtbiol.38.60865

Abstract

The Pennsylvania grotto sculpin is known from just two caves of the Nippenose Valley in central Pennsylvania, USA. They exhibit emergent troglobitic morphological traits and are the second northern-most cave adapted fish in the world. Two mitochondrial (*16S rRNA* and *D-loop* gene) and one nuclear (*S7* ribosomal protein gene intron) gene in both cave and epigean populations were sequenced. For the three markers, a large proportion of cave specimens possess unique haplotypes not found in their local surface counterparts, suggesting a vicariance in their evolutionary history. The cave population also has haplotypes from two separate lineages of surface sculpins of the *Cottus cognatus*/*bairdii* species complex. Since morphology, nuclear, and mitochondrial markers are not correlated among cave individuals, hybridization with introgression is suggested.

Keywords

Cave, Cottidae, *Cottus cognatus*, *Cottus bairdii*, hybrid, Speciation, stygobite, troglobite

Introduction

Modern biology as a science and evolutionary biology in particular have had a long history of interest in cavefishes. In evolutionary developmental biology (evo-devo), cavefishes are now viewed as a great model system (Jeffery 2001). Since the description of the first troglobitic fish, *Amblyopsis spelaea* DeKay, 1842, over two hundred species

Copyright *Luis Espinasa et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

of blind fish or fish with some degree of eye degeneration have been found in caves around the world (Behrmann-Godel 2017; Borowsky 2018).

In 2003, Espinasa and Jeffery described a previously unreported cave population of sculpins (Cottidae: Scorpaeniformes: Actinopterygii) inhabiting Eiswert #1 Cave (Stone, 1953) in the Nippenose Valley, Lycoming County, Pennsylvania. Specimens of this cave population retain some degree of pigmentation and eye functionality. They are morphologically distinct from the surface sculpins of Antes Creek located only 445 m from the cave (Fig. 1). Espinasa and Jeffery (2003) showed that the cave population differ from the surface fish by wider and more abundant mandibular pores, wider head, longer pectoral fins, and reduced eyes (Fig. 2). This population is currently the second northern-most cave adapted fish in the world (Romero and Paulson 2001; Behrmann-Godel 2017).

Eiswert #1 Cave and Antes Creek are part of the West Branch of the Susquehanna River drainage. Two closely related species of sculpins inhabit this drainage: the mottled sculpin, *Cottus bairdi* Girard, 1850, and the slimy sculpin, *Cottus cognatus* Richardson, 1836. These cottids often exhibit overlapping meristic and mensural features (Strauss 1986). Due to the occurrence of two *Cottus* species in the area and the potential for modification in taxonomic diagnostic features as a result of inhabiting the cave environment, Espinasa and Jeffery (2003) did not assign the troglomorphic sculpins to either of the species or to a new taxon. They restricted their placement to the *Cottus bairdi-cognatus* complex until further studies could be conducted. The common name given to the cave population is the Pennsylvanian grotto sculpin.

The two local surface species of cottids are known to hybridize. For example, Strauss (1986) found that at Blockhouse Creek, located in the same county as Eiswert #1 Cave, *C. bairdi* and *C. cognatus* were hybridizing. At this locality, Strauss (1986) found F1s, but no indication of F2 progeny or backcrosses. Although genetic heterozygosity in the parental samples of Blockhouse Creek tends to refute the occurrence of introgression, morphological information may support a very limited amount of introgression (Strauss 1986). Regardless of the actual degree of introgression, it appears that either the viability of hybrids is considerably reduced, or a substantial amount of intraspecific assortative mating exists. Therefore, in the surface streams of central Pennsylvania, both *C. bairdi* and *C. cognatus* gene pools remain distinct (Strauss 1986).

The purpose of this study is to establish if the Pennsylvania grotto sculpin is genetically distinct from surface sculpins and, if such is the case, from which species it derived. For this study, we sequenced two mitochondrial and one nuclear gene. The relevance of this study is not restricted to purely academic arenas. In 2002, two corporate hog farms intended to enter the Nippenose Valley (http://old.post-gazette.com/localnews/20020922farms0922p5.asp). The proposed Concentrated Animal Feeding Operations (CAFOs) were to produce over 20,000 hogs along with 5 million gallons of antibiotic and steroid rich manure each year. This surface waste could easily enter and contaminate the underground network of groundwater through the honeycombed limestone, thereby contaminating the water supply for the inhabitants of the valley. Some residents quickly organized a group called the Concerned Citizens of Nippenose

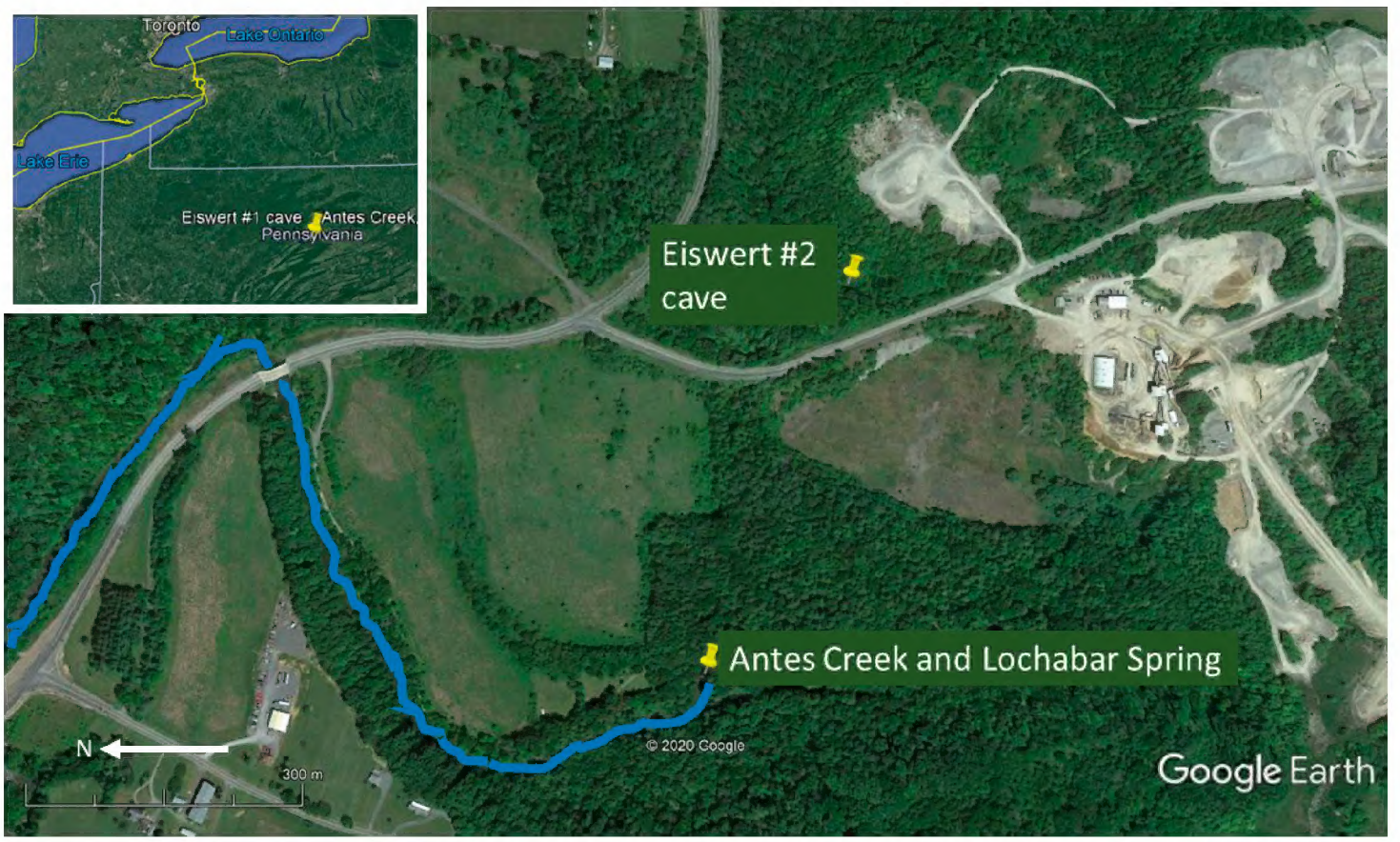


Figure 1. Nippenose Valley, showing the location of Eiswert #1 cave and Lochabar Spring, source of Antes Creek (blue). Despite being only 445 m from each other, cave and surface sculpin populations are morphologically and genetically distinct. Notice that both cave and spring are very close by to a limestone quarry, which has the potential of being a conservation threat to the Pennsylvania grotto sculpin if further developed. Inset of Pennsylvania showing location of cave. Map data 2020 Google.

Valley (CCNV). CCNV invoked a local ordinance protecting their water supply to tie up the permit applications, while mounting a comprehensive public relations campaign to build broad-based community opposition to the CAFOs. While this was occurring, the troglomorphic sculpin was found in Eiswert #1 Cave and reported in the scientific literature (Espinasa and Jeffery 2003). In particular, Espinasa and Jeffery (2003) recommended that the U.S. Fish and Wildlife Service should consider the Pennsylvania grotto sculpin to be listed under the U.S. Endangered Species Act. Afterwards, the Williamsport Sun-Gazette ran an article titled "Fish Find May Nix Pig Farms." The CCNV's efforts culminated in a victory when both CAFO applications were withdrawn. Conservation risks persist for the Pennsylvania grotto sculpins. For example, a limestone quarry is just meters away from the cave and spring (Fig. 1).

Methods

Specimens

Since the description of the Pennsylvania grotto sculpins from Eiswert #1 Cave, James C. D. Lewis (Resident Pennsylvania Fishing License number R. 703557) has identified this form in a second cave, Loose Tooth Cave, also within the Nippenose Valley,

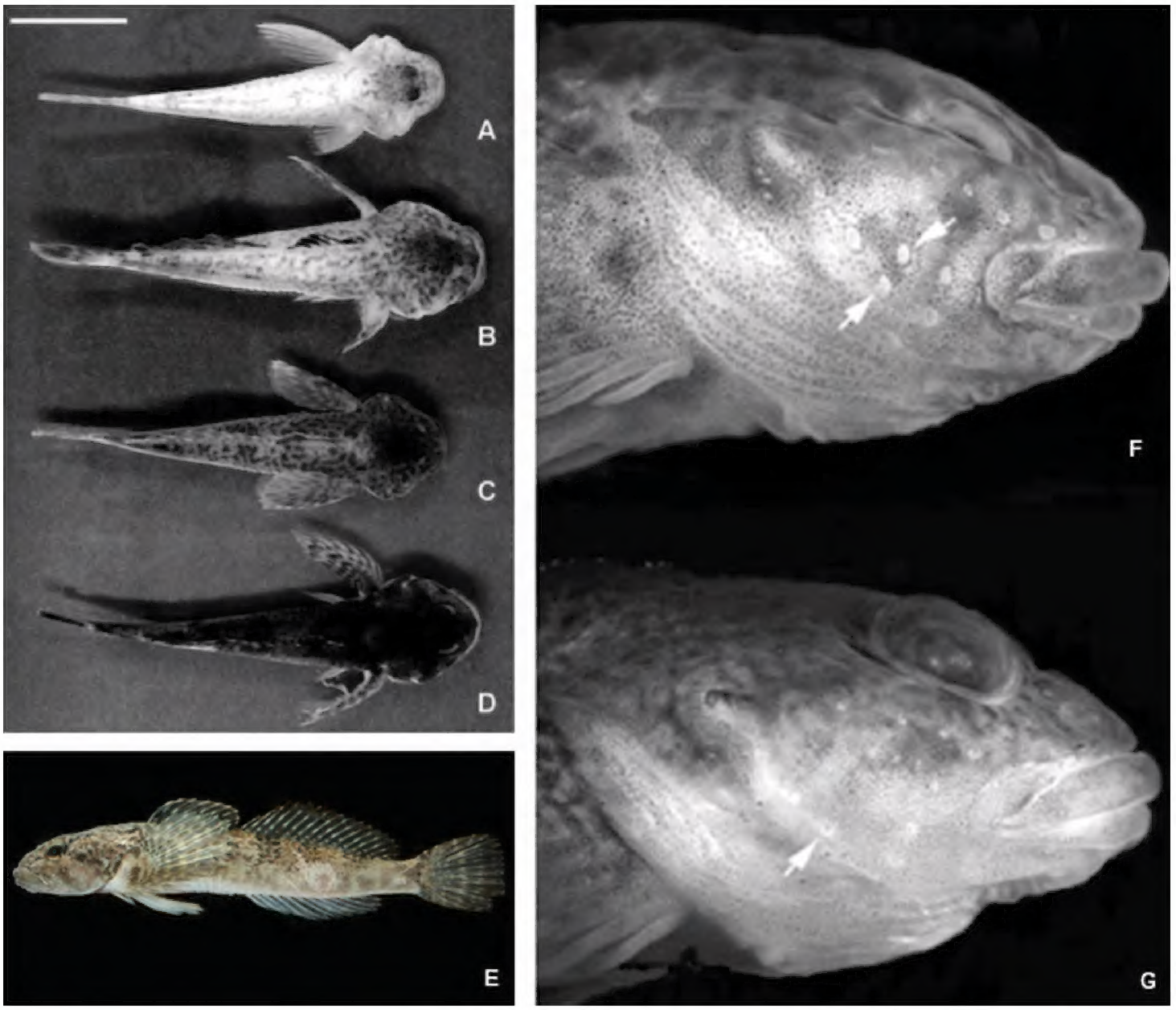


Figure 2. Nippenose Valley cave (**A–C, F**) and epigean Antes Creek (**D–E, G**) sculpins. While there is a diversity of expression of troglomorphic characters, cave fish tend to be more depigmented (**A–C** vs **D–E**), have smaller eyes (**F** vs **G**), more abundant and larger mandibular pores (**F** vs **G**, arrows point to pore III), larger heads (**B** vs **D**), and longer pectoral fins (**B** vs **D**). Scale bar: 2 cm (**A–D, F–G** modified from Espinasa and Jeffery (2003)).

2.13 km SSW of Eiswert #1. Two specimens were collected with dip nets and fin clips of fish were deposited in 100% ethanol for DNA studies. For conservation purposes, other specimens used in this study are the same as those used in Espinasa and Jeffery (2003).

Molecular data

Genomic DNA samples were obtained from:

a) Cave samples: 24 individuals from Eiswert #1 (41°9'N, 77°12'W; 8/11/02, 10/8/02, 9/16/07, 1/25/08) and two from Loose Tooth cave, Nippenose Valley, PA (41°8'N, 77°13'W; 9/16/07).

b) Surface samples: 24 individuals of presumptive *C. cognatus* from Lochabar Spring, Antes Creek, PA (41°9'28.6"N, 77°13'13.6"W; 10/8/02, 9/16/07). Lochabar

Spring is where all caves of the Nippenose Valley drain into, and it is located at a distance of 445 m WWN from the Eiswert #1 cave (Fig. 1) and 2.14 km NNE from Loose Tooth cave.

c) Surface samples: 7 individuals of *C. cognatus* from Willsey Brook, Wilsey Valley Rd. Wurtsboro, NY (41°35'N, 74°29'W; 10/7/19). According to maps provided by the New York State Department for Environmental Conservation, Willsey Brook is within a drainage inhabited by only *C. cognatus* (https://www.dec.ny.gov/animals/94615.html) and no *C. bairdii*. (Conservatiohttps://www.dec.ny.gov/animals/94617.html). Morphologic analysis of the specimens collected also showed they had three pelvic-fin rays, a diagnostic feature of *C. cognatus*, while *C. bairdii* has four pelvic-fin rays, thus confirming specimens belong to *C. cognatus*.

Standard methods for DNA purification were followed. Total DNA was extracted from tail clippings using Qiagen's DNEasy Tissue Kit. Two mitochondrial and one nuclear marker were amplified and sequenced, each as a single fragment. The primers used were *16S*ar and *16S*b primer pair for *16S rRNA* (Edgecombe et al. 2002), DL1F and DL4R primer pair for *D-loop* (Bagley and Gall 1998), and S7RPEX1F and S7RPEX2R primer pair for *S7* ribosomal protein gene first intron (Chow and Hazama 1998). Amplification was carried out in a 50 µl volume reaction, with 1.25 units of AmpliTaq DNA Polymerase (Perkin Elmer, Foster City, CA, USA), 200 µM of dNTPs, and 1 µM of each primer. The PCR program amplification cycles as in Edgecombe et al. (2002). PCR amplified samples were purified with Qiagen's QIAquick Gel Extraction Kit and directly sequenced using an automated ABI Prism 3700 DNA analyzer as in Espinasa et al. (2007). Chromatograms obtained from the automated sequencer were read and analyzed with the program Sequencher 3.0. Primers and 5' or 3' short end fragments of sequences not shared by all individuals in the analyses were trimmed. Sequences were compared against highly similar sequences available in GenBank using BLASTn. Sequences were aligned with CLUSTALW2 (Higgins et al. 1992; Thompson et al. 1994). A minimum haplotype network was obtained using PopArt program (Leigh and Bryant 2015). In addition, a maximum likelihood tree was inferred using PhyML v20160115 ran with the following model and parameters: -0 tir –pinv e –nclasses 4 –bootstrap 10 f m –alpha e, as implemented in CLUSTALW2. Branch supports were computed out of 100 bootstrapped trees. The analyses also included the sequences most similar to the Nippenose cave cottids found in GenBank through BLAST. Nei's gene diversity (h), Nei's population differentiation (*Gst*), and gene flow (Nm) were estimated with software POPGENE version 1.32 (Yeh 1999).

To determine if specimens harboring either of the mitochondrial haplotype lineages (see below) belonged to two separate species or to a single reproductive population that had undergone hybridization and introgression, morphology, nuclear, and mitochondrial markers were compared to evaluate if they were linked or correlated among individuals of the cave population. The assumption being that if they were two different species, specific nuclear haplotypes and morphologies would be found only in individuals carrying a specific mitochondrial haplotype.

Morphology

Analyses included data from the specimens described by Espinasa and Jeffery (2003), in addition to six new specimens from Eiswert #1 Cave and two from Loose Tooth Cave collected on 16 September 2007 for a total of 23 specimens. Measurements of eye and mandibular pore #3 lengths were obtained using a Motic K series stereomicroscope with a fitted eyepiece micrometer. A dial caliper was used to measure to the nearest 0.1 mm the standard length of the pectoral fin length and the head width. Eye, pore #3, pectoral fin, and head width were plotted against standard length and linear regressions were obtained using Microsoft Excel. From these plots, each of the specimens was assigned a numerical value to indicate if they had a troglobitic (0), intermediate (1), or surface-like (2) appearance based on their location with respect to the linear regression line. To do this, the seven individuals furthest above, nine closest, and the seven individuals furthest below the linear regression line were identified (Fig. 3). Overall troglomorphic appearance of an individual within the population was calculated by adding the numbers for the four characters. If an individual, when in comparison to the linear regression of the population, had small eyes and a large pore, large fin, and large head, the lowest total value it could receive would be "0". Whereas for an individual with large eyes and a small pore, fin, and head, the highest value it could receive would be "8". A non-parametric Mann-Whitney U test for ordinal data was conducted to test if specimens harboring either of the mitochondrial haplotype lineages were equally troglobitic in morphologic appearance.

Results

Molecular data

The two caves as a population had a haplotype composition completely different from the local surface population in Antes Creek. While the surface population has haplotypes corresponding to *C. cognatus*, the cave population has haplotypes shared with two sculpin lineages: *C. cognatus* and *C. bairdi*.

For the mitochondrial *16S rRNA*, analyses were performed on a 526 bp sequence fragment shared by all individuals. A total of six haplotypes were found: *Cottus* sp. 'Nippenose Valley' haplotypes 1 and 2 (GenBank accession nos. GQ280792 and GQ280792) and *Cottus* sp. 'Nippenose Valley' haplotypes A (GQ267192), B (GQ267193), C (GQ267194), and D (GQ267190). Both maximum likelihood and haplotype network analyses of the *16S rRNA* haplotypes of the cave and surface populations identified two clearly distinct clades (Fig. 4). The first clade includes only cave specimens with haplotypes 1 and 2, and the second clade includes both surface fish from Antes Creek and cave fish with haplotypes A–D. For Loose Tooth Cave specimens, one individual had haplotype A and the other haplotype 2, which were the two most common haplotypes in Eiswert #1 Cave.

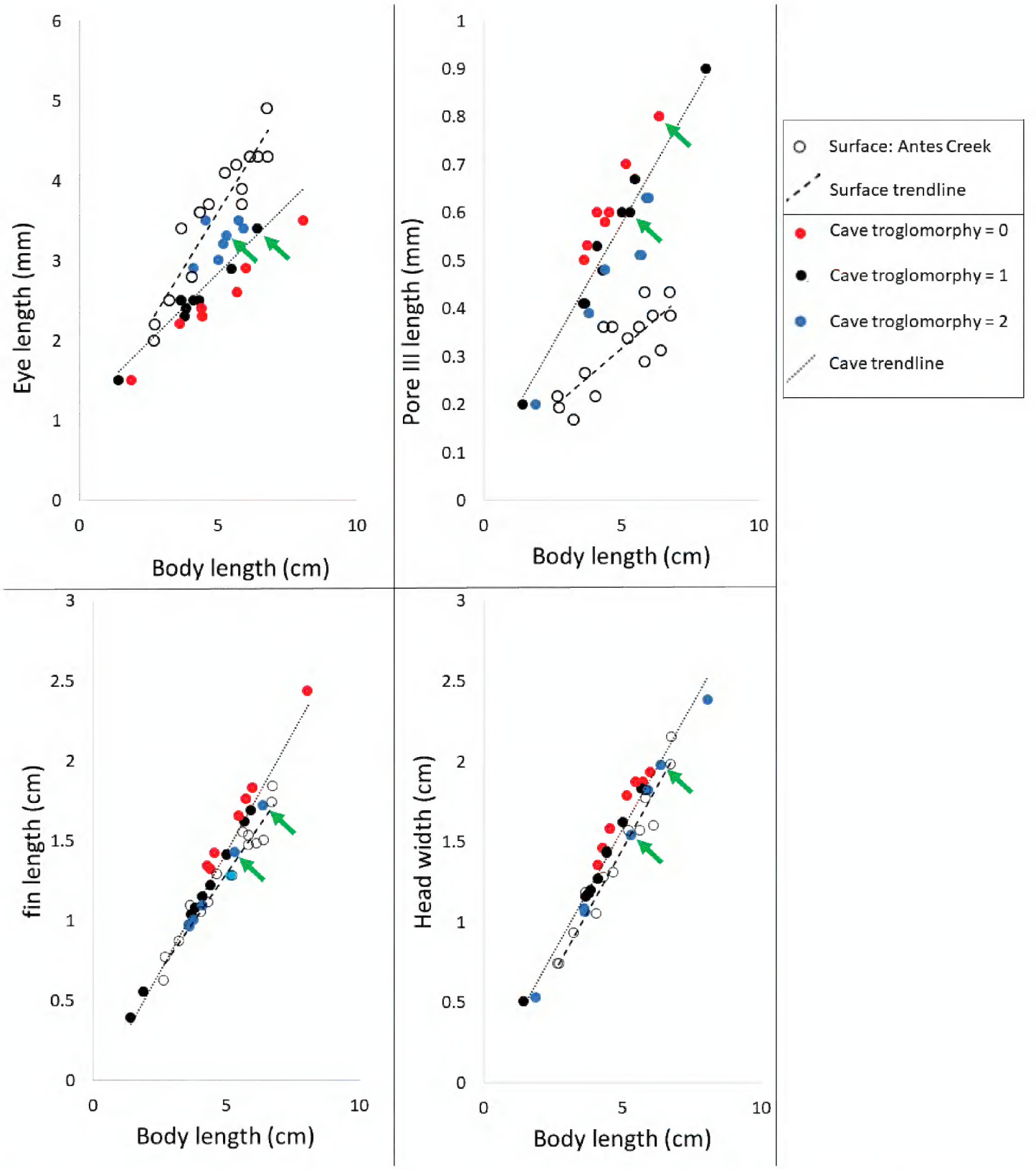


Figure 3. Morphometric comparisons between cave (closed circles) and epigean Antes Creek (open circles) sculpins (Modified from Espinasa and Jeffery 2003). Green arrows point to Lose tooth cave specimens, showing they are within the morphologically range of Eiswert #1 cave specimens. From these plots, cave specimens were further evaluated to assign their overall individual troglomorphic appearance. The seven individuals furthest above, nine closest, and the seven individuals furthest below the linear regression line were identified. Each were correspondingly assigned a numerical value to indicate they had a troglobitic (0; red), intermediate (1; black), or surface-like (2; blue) appearance.

Mean sequence divergence (*p*-distance in parenthesis) of the *16S rRNA* among the two clades was 11 (range 10–12; 1.9–2.2%) substitutions. Mean intra-clade sequence divergence within clades was of 1 (maximum of 2; 0.3% substitutions). Sequences in the first clade were most similar to *C. bairdii* (GenBank accession no. AY539018)

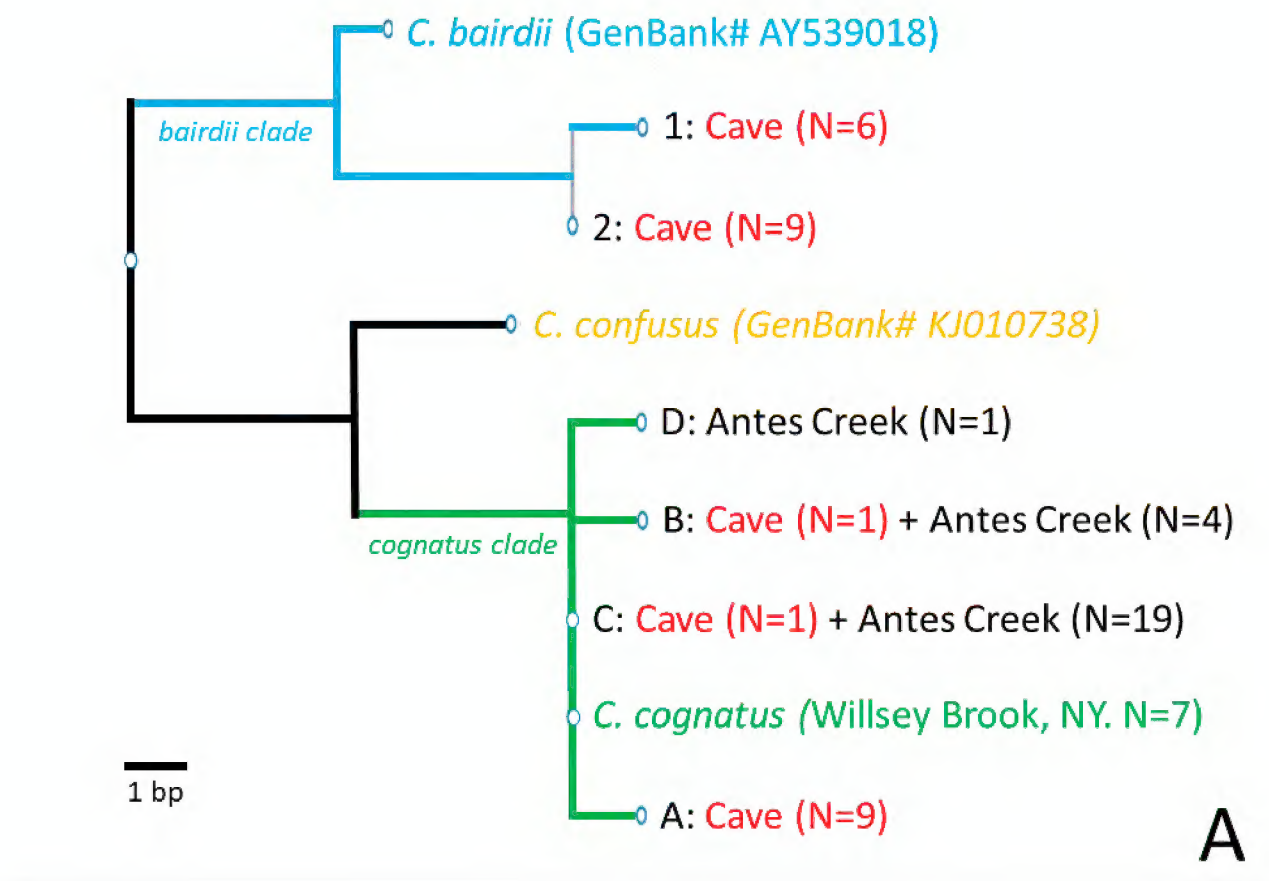


Figure 4. A maximum likelihood tree of the *16S rRNA* haplotypes found in cave and surface specimens. **B** minimum spanning haplotype network. Most similar sequences obtained in GenBank through BLAST analyses plus a sequence of *C. cognatus* from Willsey Brook, NY, are also included in the trees. Two distinct lineages were identified; one that includes *C. bairdii* specimens and one that includes *C. cognatus*. Cave haplotypes within the *C. bairdii* clade are identified by numerals (1–2) and Cave and Antes Creek haplotypes within the *C. cognatus* clade are identified by letters (**A–D**). Notice that cave individuals are found within both lineages. Despite being only 445 m apart, some haplotypes are present exclusively in cave individuals and not found in the Antes Spring population. Loose Tooth cave specimens had haplotype 2 and A.

whose sequence divergence was between 4–5 (0.7–0.9%) substitutions, and a group of three identical sequences (MT539220–MT539222). These last sequences were not included in the analyses because they were labelled by the authors as both *C. bairdii* and *C. cognatus*, despite having identical sequence, and thus are of doubtful provenance. Sequences in the second clade were most similar to seven *C. cognatus* from Willsey Brook, New York, that were sequenced in this study. Median sequence divergence between these *C. cognatus* and the Pennsylvania cave and surface members of the clade was 0 (range 0–2; 0–0.3%) substitutions. The next most similar sequence found with BLAST analyses was a *C. confusus* (KJ010738), with a sequence divergence of 5–6 (0.9–1.1%) substitutions.

16S haplotypes for all 24 surface specimens from Lochabar Spring on Antes Creek belonged to the second clade, which included *C. cognatus*. The fact that the Antes Creek population is mainly inhabited by *C. cognatus* is further supported by 90.4% of the cottids examined from Antes Creek possessing three pelvic rays (Espinasa and Jeffery 2003), which is a diagnostic character of *C. cognatus*.

Sequence data of the *16S rRNA* of the Pennsylvania grotto sculpin shows that, despite inhabiting a cave that is only 445 m from the aforementioned Lochabar Spring, it has a haplotype composition completely different from the local surface population in Antes Creek. Individuals from the cave harbor haplotypes from the first *C. bairdii* clade and from the second *C. cognatus* clade (Fig. 4). Of the 26 specimens analyzed, 15 (57.6%) had a haplotype within the *C. bairdii* clade and 11 (42.3%) within the *C. cognatus* clade. The cave population had a larger number of haplotypes (5) and a much higher gene diversity (h=0.720) than the Antes Creek population (3; h = 0.348). Haplotypes were more equitably distributed as the most common haplotype was in only 34.6% of the cave population, while in the Antes Creek, the most common haplotype was present in 79.1% of the surface population. Based on the *16S rRNA* data, there is strong population differentiation between the cave and surface populations (*Gst* = 0.361) with intermediate level of gene flow among them (Nm = 0.883).

Evidence suggests that the cave population is not simply surface *C. cognatus* from Antes Creek that happened to swim inside the cave and share the environment with local *C. bairdii*. As mentioned previously, of the 26 cave samples, 11 (42.3%) had haplotypes within the *C. cognatus* clade. Of these, 8 (30.7%) had a unique *C. cognatus* haplotype not found in any of the 24 surface Antes Creek *C. cognatus* specimens analyzed (Fig. 4). Thus, even within the individuals harboring haplotypes within the *C. cognatus* clade, the genetic structure is different between the cave and the surface populations.

Phylogenetic analysis of a second mitochondrial marker, the *D-loop* (Fig. 5), corroborated the two distinct clades identified with the *16S rRNA* data. A 786-bp fragment was analyzed in 15 cave and 11 surface individuals from Antes Creek. Twenty haplotypes were found (GenBank accession nos. GQ290440–GQ290447 in the first clade and GQ290448–GQ290459 in the second clade). Median sequence divergence between the two clades was 47 (range 46–49; 5.8%–6.2%) substitutions. Median sequence divergence within clades was of 1 (maximum 3; 0.3%) substitutions. BLAST analyses showed that the most similar sequence to the first clade in GenBank were two

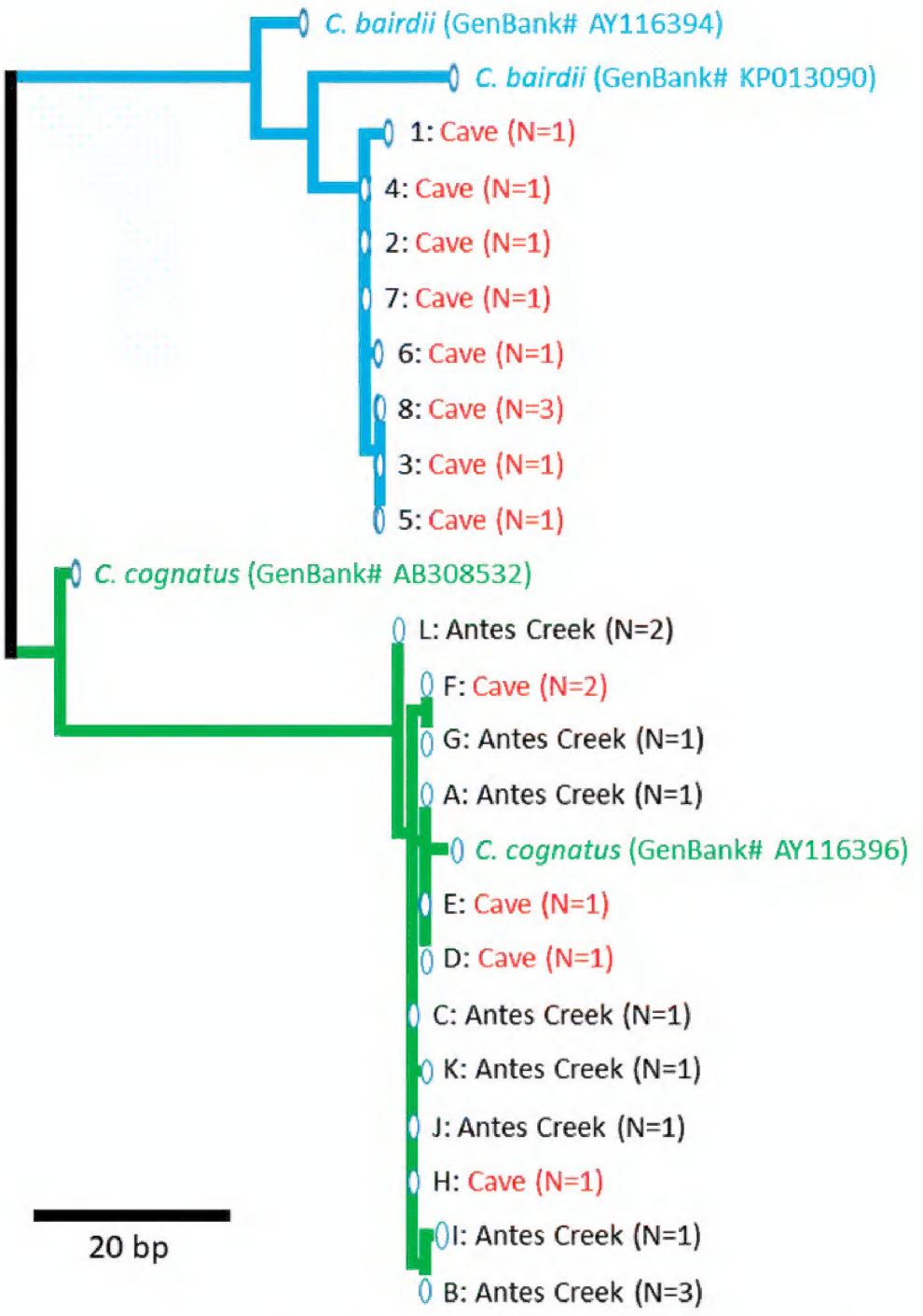


Figure 5. Maximum likelihood tree of the *D-loop* haplotypes found in cave and surface specimens. Most similar sequences obtained in GenBank through BLAST analyses are included in the tree. In agreement with the *16S rRNA* data, two distinct lineages were identified; one that includes *C. bairdii* specimens and one that includes *C. cognatus*. Cave haplotypes within the *C. bairdii* clade are identified by numerals (1–8) and Cave and Antes Creek haplotypes within the *C. cognatus* clade are identified by letters (A–L). Notice that cave individuals are found within both lineages.

sequences ascribed to *C. bairdii* (KP013090 and AY116394). Their sequence divergence against the cave specimens was between 21–22 (2.6–2.7%) substitutions. The second clade had as its most similar sequence in GenBank two sequences ascribed to *C. cognatus* (AY116396 and AB308532). Sequence divergence between these *C. cognatus* and the Pennsylvania cave and surface members of the clade ranged from 8–11 (1.0–1.3%) substitutions for AY116396 and 32–35 (4.0–4.4%) for AB308532. Just like with the *16S rRNA*, all 15 Antes Creek surface specimens analyzed harbored a haplotype within the *C. cognatus* clade, while those from the cave were mixed (Fig. 5).

Five individuals had a *D-loop* haplotype within the *C. cognatus* clade and eight within the *C. bairdii* clade. For the cave specimens harboring a mitochondrial *D-loop* haplotype within the *C. cognatus* clade, the haplotypes themselves were also distinct from the haplotypes hosted by *C. cognatus* surface fish (Fig. 5). As was expected from mitochondrial linked markers, *D-loop* results corroborate the previous findings with the *16S rRNA* that the genetic structure of the cave population is not simply the result of joining surface *C. cognatus* from Antes Creek that happened to swim inside the cave and share the environment with local *C. bairdii*.

Results from the *16S rRNA* and *D-loop* suggest that the cave population has haplotypes shared with two distinct lineages of surface sculpins. Since these two markers are mitochondrial, maternally inherited and linked, they cannot fully resolve if there has been introgression of two phyletic lines within the cave population. In order to resolve this, a nuclear marker (the *S7*) was sequenced. A 526-bp fragment of the nuclear S7 locus from 19 cave specimens showed the presence of two haplotypes differing by one bp (GenBank accession nos. MW039591 and MW039592). One of these haplotypes was identical to the sequence obtained from all 15 surface fish sequenced from Antes Creek and from the seven surface *C. cognatus* from Willsey Brook, New York. BLAST analysis of both haplotypes showed as most similar a *C. microstomus* (KY246946), from which it differs by 7–8 bp (1.3–1.5%), but it was noticed that no *S7* gene sequence for *C. bairdii* or *C. cognatus* has yet been uploaded to GenBank. Despite inhabiting a cave that is only 445 m from Antes Creek, it has an *S7* haplotype composition completely different from the local surface population in Antes Creek. While all the surface Antes Creek fish (N = 15) had the *C. cognatus* haplotype as homozygous (G base), 15 out of 19 (78.9%) cave fish had a haplotype not found in the surface fish, either as homozygous (A base) or heterozygous (A/G bases). Of the 19 cave specimens sequenced, four had the *C. cognatus* haplotype, five had the other haplotype, and ten were heterozygous for both haplotypes (Fig. 6). Loose Tooth cave specimens were homozygous, one for the A base and the other for the G base.

Of the 11 cave specimens analyzed for *S7* that harbor the *C. bairdi* clade mitochondrial haplotypes, two of them had the *S7* found in *C. cognatus*, and the rest were either heterozygous or had the other haplotype. Likewise, of the eight cave specimens that harbored the *C. cognatus* clade mitochondrial haplotypes, only two had exclusively the *C. cognatus S7* haplotype. The rest were either heterozygous or had the other haplotype (Fig. 6).

Morphology

As described by Espinasa and Jeffery (2003), the cave specimens from both caves were paler in color than the surface specimens, had smaller eyes, wider and more abundant mandibular pores, a wider head, and longer pectoral fins (Fig. 2). Some individuals after being exposed to light darken somewhat in color.

The range of troglomorphic appearance spanned from "0" when all characters analyzed in an individual were troglomorphic, to "8" when they were all epigeomorphic.

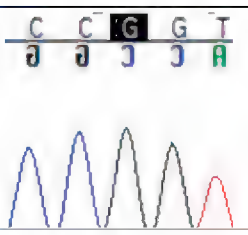
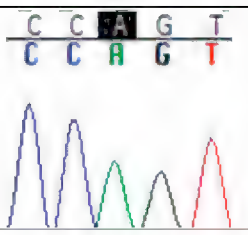
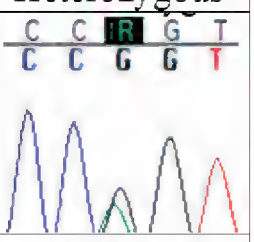
	G	A	Heterozygous
			
C. cognatus Willsey Brook, NY	N=7	-	-
C. cognatus Surface Antes Creek, PA	N=15	-	-
Cave with *C. cognatus* mitochondrial haplotypes	N=2	N=2	N=4
Cave with *C. bairdii* mitochondrial haplotypes	N=2	N=3	N=6

Figure 6. Two haplotypes were found for the *S7* ribosomal protein gene intron, which differed by a G/A base. Some individuals were heterozygous, as evidenced by a double bump in the chromatogram. *C. cognatus* individuals from Antes Creek or from New York had exclusively the *S7* haplotype with the G. Cave specimens with *C. cognatus* mitochondrial haplotypes could be homozygous or heterozygous for both *S7* haplotypes. Likewise, cave specimens with *C. bairdii* mitochondrial haplotypes could also be homozygous or heterozygous for both *S7* haplotypes. Thus, there is no correlation between nuclear haplotypes and mitochondrial clades, suggesting introgression within the cave population as a single reproductive unit. The Loose Tooth cave specimen was one of the specimens that had *C. bairdii* haplotype and an A, and the other specimen had *C. cognatus* haplotype and a G.

When a value for troglobitic appearance was assigned to each cave individual based on four distinct characters, it was found that troglobitic characteristics were spread arbitrarily among each group of mitochondrial or nuclear haplotypes, with no clear distinction between the two groups. Cave individuals harboring the *C. bairdii* clade mitochondrial haplotype had a mean troglomorphic index of 4.14 ($\tilde{x}$ = 4, stdev = 1.29, Range = 2–7, N = 15) and the cave individuals harboring the *C. cognatus* mitochondrial haplotype had a mean troglomorphic index of 3.77 ($\tilde{x}$ =4, stdev = 1.71, Range = 1–7, N = 11), which was not different from individuals belonging to the *C. bairdii* clade (P = 0.509). Furthermore, specimens from Loose tooth had morphologies within the ranges of Eiswert #1cave specimens (Fig. 3), suggesting they are part of the same population.

We also found morphological evidence of an independent assortment of morphological characters without regard to their mitochondrial haplotypes, as is expected in a reproductive population where unlinked genes/alleles segregate independently of each other. One of the clearest examples is with two specimens that have a *C. bairdii* clade mitochondrial haplotype. One specimen has a pore III width that is surface-like (Troglomorphy=2), but eye size is troglobitic-like (Troglomorphy=0). On the other hand, the other specimen has the exact opposite combination; pore III width troglobitic-like (Troglomorphy = 0) and eye size is surface-like (Troglomorphy = 2). Similar examples are also found within specimens harboring the *C. cognatus* clade mitochondrial haplotype.

Discussion

Eiswert #1 and Loose Tooth caves are both located within the same Nippenose karst valley and only 2.13 km from each other. Hydrologically, they are most likely part of the same subterranean drainage. Morphologically (Fig. 3), specimens of both populations are similar and the two Loose Tooth cave specimens have the most common *16S rRNA* and *S7* haplotypes found in Eiwert #1 cave population. The geographical proximity as well as morphological and genetic data support that specimens from both caves are part of the same underground population and it is likely that in the rainy season, when floods occur, specimens can swim from one cave to the other.

Eiswert #1 Cave is only 445 m away from the Lochabar Spring and the surface Antes Creek (Fig. 1). The cavefish have variable levels of pigmentation (Fig. 2) and at least some cave individuals have functional vision (Espinasa and Jeffery 2003). Could it be that the Pennsylvania grotto sculpin are nothing more than *C. cognatus* surface fish from Antes Creek that happened to swim upstream into the cave and became pale in the process? Our results reject this notion. A large proportion of individuals from the cave population have unique mitochondrial and nuclear haplotypes not found in the local surface counterparts, suggesting isolation in their evolutionary history. Twenty-four surface specimens were studied for the *16S rRNA* from the surface locality of Antes Creek. All individuals exhibited one of three haplotypes within a clade that includes *C. cognatus*. On the contrary, of the 26 cave specimens analyzed, only two individuals (7.6%) had the same haplotype as the surface population (Fig. 4). This result was supported with *D-loop* analyses of 12 cave and 15 surface individuals from Antes Creek, where not a single cave specimen shared a specific haplotype with the local surface population (Fig. 5). Likewise, 15 cave individuals out of 19 had a nuclear *S7* haplotype not found in the local surface population (Fig. 6).

Our results show that the cave population possesses mitochondrial *16S rRNA* and *D-loop* haplotypes shared with two distinct lineages of surface-dwelling sculpins. One line includes specimens assigned to *C. cognatus* and the other line to specimens assigned to *C. bairdi*. What could account for this? One scenario is that surface specimens from both species are simply entering the cave and the single cave population is a mixture of two surface species. Our results do not support this scenario. If two species were present, cave individuals harboring the mitochondrial haplotypes of the *C. cognatus* and the *C. bairdii* clade would have distinct and correlated nuclear and morphologic characters. Instead, we found that morphology, nuclear, and mitochondrial markers were unlinked among individuals of the cave population, suggesting a single reproductive population with introgression. Both nuclear haplotypes are found in cavefish that harbor either of the mitochondrial haplotype clades. Likewise, the nuclear *S7* shows heterozygosity in some individuals. Furthermore, there is evidence of an independent assortment of morphological characters without regard to their mitochondrial or nuclear haplotypes, as is expected in a reproductive population where unlinked gene/alleles segregate independently of each other.

If the Pennsylvania grotto sculpin as a population have a different genetic structure from the local surface *C. cognatus* and *C. bairdii*, what is their origin? One possibil-

ity is that the cave population derives from a single species who just happens to host divergent, ancient mitochondrial haplotypes. Coalescent theory shows that sometimes gene variants sampled from within a population may have originated from a common ancestor that antedates the split of its own species (Fumey et al. 2018). In essence, the presence of mitochondrial DNA haplogroups found in two different species does not necessarily imply that a population is composed of two separate species, i.e., the time of separation of the populations is not necessarily equal, not even close, to the time of coalescence of the mtDNA sequences (Fumey et al. 2018). The other scenario is that there was a hybridization event in the past between different members of the *C. cognatus*/*bairdi* species complex, whose vestiges comprise the cave population.

While the distinction may not be fully resolved until genomic studies are performed, it is our hypothesis that the hybridization scenario is the most likely. The number of rays in the pelvic fin is used as a diagnostic character to differentiate species of *Cottus* in this region. *Cottus cognatus* has three pelvic fin rays, while *C. bairdii* has four. If the Pennsylvania grotto sculpin derived exclusively from either species, it would be expected that the cavefish would conform morphologically to the species from which it derived and would either have three or four rays. Espinasa and Jeffery (2003) showed that the cave population does not fully ascribe to either. Eighty percent (80%) of the cave population has three rays, 6% has four rays, but, most significantly, 13% are asymmetric in which on one side they have three rays and on the other side they have four. Bilateral asymmetry has been employed as a criterion of developmental instability on the assumption that coordination among loci is disrupted by hybridization (Strauss 1986). In the Espinasa and Jeffery (2003) study, they showed that the percentage of asymmetric individuals for troglomorphic fish is an order of magnitude higher than in *C. bairdi* and *C. cognatus* but is comparable to the bilateral asymmetry shown in hybrid individuals found at Blockhouse Creek (Strauss 1986).

Hybridization between surface *C. bairdi* and *C. cognatus* has been reported at Blockhouse Creek (Strauss 1986), which is in the same county as Eiswert #1 Cave. Unlike the cave population, there is no indication of F2 progeny or backcrosses at this surface locality. What could account for this? A conjectural idea is that the cave habitat provided an adaptive peak where hybrids could survive that is normally not available within the adaptive landscape of surface sculpins. In other words, hybrids in surface habitats are probably eliminated by competitive exclusion of the non-hybrid surface species. The cave somehow provided a safe haven for a hybrid population, and now they may be undergoing speciation by isolation, as evidenced by the presence of unique haplotypes not found in local surface populations. While we do not discount that currently there may be surface fish that manage to enter the cave and even successfully hybridize with the cave population, the *16S rRNA* data shows there is strong population differentiation between the cave and surface populations (*Gst* = 0.361) with intermediate level of gene flow among them (Nm = 0.883). Another line of evidence in support of past hybridization when the cave population first evolved is the high levels of diversity. When two populations mix, hybrids acquire haplotypes from both parental populations. The cave population has a larger number of *16S rRNA*

haplotypes (n = 5) and a much higher gene diversity (h = 0.720) than the Antes Creek population (n = 3; h = 0.348).

Our results are not the first to report speciation by way of hybridization in cottids. In 2005, Nolte et al. (2005) described an invasive new lineage of sculpins in the Rhine River system of Central Europe. It was also found with mitochondrial haplotypes that the invasive sculpins are in fact hybrids between two species of the *Cottus gobio* complex. This invasive hybrid population of sculpins was determined to possess a unique ecological potential that does not occur in either of the source populations, which allowed for the colonization of new habitats that previously lacked sculpins. They concluded that hybridization led to a new, adaptationally distinct sculpin lineage.

Our results are similar to those of Nolte et al. (2005), with the difference that instead of a portion of the Rhine River, a cave was colonized. We also hypothesize that the cave population has been relatively isolated from both surface species, allowing for divergent and adaptive evolution to the cave environment. Evidence of this divergence is that most cavefish exhibit haplotypes not found in any of the surface specimens studied. If significant levels of introgression were present between the surface and cave populations, this degree of haplotype distinctiveness should not be found. Furthermore, the cave population has a unique and distinctive morphology, most likely adaptive to the cave environment, since enlarged fins and mandibular pores suggest an increased sensory capacity while in complete darkness.

Conclusions

Molecular and morphological data support the hypothesis that the Pennsylvania grotto sculpin is a distinct sculpin from *C. cognatus* and *C. bairdi*. Furthermore, data suggest that the cave population's evolutionary history may include an ancestral hybridization event between the separate members of the *C. cognatus/bairdii* species complex, but that currently there is limited gene flow from surface *Cottus* populations into the cave population. Such isolation accounts for the cave population's genetic and morphologic uniqueness. Based on these results, it is proposed that the Pennsylvania grotto sculpin deserves recognition as an independent species taxon from *C. bairdii* and *C. cognatus*. Recognition as an independent species will also help support current conservation efforts.

Acknowledgements

We thank all of the concerned citizens of the Nippenose Valley, especially David Hollick, for all of their encouragement, support, and friendship, without which this study could not have been done. We also would like to thank James C. D. Lewis for collecting the samples. Katherine D. Amodeo performed the morphological measurements. The students of the BIOL493: Molecular Biology Fall 2007 and BIOL320: Genetics Fall 2018 courses at Marist College performed some of the molecular techniques. Pa-

tricia Ornelas Garcia helped with developing the haplotype networks. We also thank Monika Espinasa, Jack Wimmershoff, Shivani Patel and Marylena Mesquita for proofreading and English language editing. Partial support for this study from the School of Science at Marist College.

References

Bagley MJ, Gall GAE (1998) Mitochondrial and nuclear DNA sequence variability among populations of rainbow trout (*Oncorhynchus mykiss*). Molecular Ecology 7(8): 94–961. https://doi.org/10.1046/j.1365-294x.1998.00413.x

Borowsky, R (2018) Cavefishes. Current Biology 28(2): R60–R64. https://doi.org/10.1016/j.cub.2017.12.011

Chow S, Hazama K (1998) Universal PCR primers for *S7* ribosomal protein gene introns in fish. Molecular Ecology 7(9): 1255–1256. https://doi.org/10.1007/s40071-016-0122-5

De Kay JE (1842) Zoology of New-York: Or, The New York Fauna; Comprising Detailed Descriptions of All the Animals Hitherto Observed Within the State of New York, with Brief Notices of Those Occasionally Found Near Its Borders, and Accompanied by Appropriate Illustrations (Vol. 1). Carroll and Cook, printers to the Assembly, 558 pp. https://doi.org/10.5962/bhl.title.13735

Edgecombe GD, Giribet G, Wheeler WC (2002) Phylogeny of Henicopidae (Chilopoda: Lithobiomorpha): a combined analysis of morphology and five molecular loci. Systematic Entomology 27: 31–64. https://doi.org/10.1046/j.0307-6970.2001.00163.x

Espinasa L, Jeffery W (2003) A Troglomorphic Sculpin (Pisces: Cottidae) Population: Geography, Morphology and Conservation Status. Journal of Cave and Karst Studies 65(2): 93–100. https://caves.org/pub/journal/PDF/V65/v65n2-Espinasa.pdf

Fumey J, Hinaux H, Noirot C, Thermes C, Rétaux S, Casane D (2018) Evidence for late Pleistocene origin of *Astyanax mexicanus* cavefish. BMC Evolutionary Biology 18(43): 1–19. https://doi.org/10.1186/s12862-018-1156-7

Girard CF (1850) A monograph of the freshwater *Cottus* of North America. Proceedings of AAAS (2nd meeting, 1849): 409–411. https://doi.org/10.1186/s12862-018-1156-7

Higgins DG, Bleasby AJ, Fuchs R (1992) CLUSTAL V: improved software for multiple sequence alignment. Computer Applications in Bioscience 8(2): 189–191. https://doi.org/10.1093/bioinformatics/8.2.189

Leigh, JW, Bryant D (2015) PopART: Full-feature software for haplotype network construction. Methods in Ecology and Evolution 6(9): 1110–1116. https://doi.org/10.1111/2041-210X.12410

Nolte AW, Freyhof J, Stemshorn KC, Tautz D (2005) An invasive lineage of sculpins, *Cottus* sp. (Pisces, Teleostei) in the Rhine with new habitat adaptations has ogriginated from hybridization between old phylogeographic roups. Proceedings of the Royal Society B: Biological Sciences 272(1579): 2379–2387. https://doi.org/10.1098/rspb.2005.3231

Richardson J, Swainson W, Kirby W (2012) Fauna Boreali-Americana; Or, The Zoology of the Northern Parts of British America: Containing Descriptions of the Objects of Natural

History Collected on the Late Northern Land Expeditions Under Command of Captain Sir John Franklin, RN (Vol. 2). Cambridge University Press.

Stone RW (1953) Caves of Pennsylvania: The American Caver. Bulletin of the National Speleological Society 15: 112–113. https://caves.org/pub/journal/NSS%20Bulletin/Vol_15.pdf

Strauss RE (1986) Natural Hybrids of the Freshwater Sculpins *Cottus bairdi* and *Cottus cognatus* (Pisces: Cottidae): Electrophoretic and Morphometric Evidence. American Midland Naturalist 115(1): 87–105. https://doi.org/10.2307/2425839

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research 22(22): 4673–4680. https://doi.org/10.1093/nar/22.22.4673

Yeh FC (1999) POPGENE, v.1.31. http://www.ualberta.ca/~fyeh/fyeh